

Statistica della Formazione

Slides 2

A.A. 2020-2021

Docente: ANNA LINA SARRA

Modulo 1: elementi di statistica descrittiva

- **Distribuzioni statistiche**
- **Rappresentazioni grafiche**

Distribuzioni statistiche

- L'effetto dell'operazione di determinazione della modalità con cui ognuno dei caratteri si presenta in ciascuna unità del collettivo determina la “**distribuzione**” del collettivo secondo i caratteri considerati.
- La distribuzione indica come le modalità dei caratteri si distribuiscono nelle unità del collettivo.

Distribuzioni disaggregate

| ID | GENERE | TIPO DI DISTURBO |
|----|---------|------------------|
| 2 | MASCHIO | DISLESSIA |
| 3 | MASCHIO | DISGRAFIA |
| 7 | FEMMINA | DISCALCULIA |
| 9 | FEMMINA | DISCALCULIA |
| 10 | FEMMINA | DISGRAFIA |
| 12 | FEMMINA | DISGRAFIA |
| 15 | FEMMINA | DISLESSIA |
| 16 | FEMMINA | DISGRAFIA |
| 18 | FEMMINA | DISGRAFIA |
| 19 | FEMMINA | DISCALCULIA |

Distribuzioni di frequenza

| | numero |
|-------------|--------|
| MASCHIO | 2 |
| FEMMINA | 8 |
| | |
| DISLESSIA | 2 |
| DISGRAFIA | 5 |
| DISCALCULIA | 3 |

Distribuzioni statistiche disaggregate

Si consideri un collettivo statistico di N unità, dove si sia osservato il carattere X .

Si chiama **distribuzione statistica disaggregata** secondo il carattere X l'insieme delle osservazioni (rappresentate da numeri o da espressioni verbali) relative alle N unità del collettivo.

In simboli, la distribuzione disaggregata sarà indicata come x_1, x_2, \dots, x_N dove x_1 è l'osservazione relativa all'unità identificata dal numero 1, x_2 l'osservazione relativa all'unità identificata dal numero 2 e così via.

Esempi di distribuzioni statistiche disaggregate

| SOGGETTO | GENERE | TIPO DI DISTURBO |
|----------|---------|------------------|
| 1 | MASCHIO | DISGRAFIA |
| 2 | MASCHIO | DISCALCULIA |
| 3 | MASCHIO | DISGRAFIA |
| 4 | MASCHIO | DISLESSIA |
| 5 | FEMMINA | DISLESSIA |
| 6 | FEMMINA | DISLESSIA |
| 7 | MASCHIO | DISGRAFIA |
| 8 | MASCHIO | DISCALCULIA |
| 9 | MASCHIO | DISCALCULIA |
| 10 | FEMMINA | DISGRAFIA |
| 11 | MASCHIO | DISCALCULIA |
| 12 | MASCHIO | DISCALCULIA |
| 13 | FEMMINA | DISLESSIA |
| 14 | FEMMINA | DISLESSIA |
| 15 | MASCHIO | DISCALCULIA |
| 16 | MASCHIO | DISLESSIA |
| 17 | MASCHIO | DISCALCULIA |
| 18 | FEMMINA | DISLESSIA |
| 19 | MASCHIO | DISCALCULIA |
| 20 | FEMMINA | DISLESSIA |

Distribuzioni di frequenze

- L'operazione di raggruppamento delle unità statistiche viene realizzata mediante la **classificazione** o lo **spoglio dei dati**.

Per **frequenza** si intende il numero di volte che una data modalità si presenta nel collettivo statistico.

Si chiama **distribuzione di frequenze** lo schema con cui si associa a ciascuna modalità del carattere X la rispettiva frequenza.

| Modalità (x_i) | Frequenza (n_i) |
|------------------------------------|-------------------------------------|
| x_1 | n_1 |
| x_2 | n_2 |
| \vdots | \vdots |
| x_k | n_k |
| Totale | N |

$$\sum_{i=1}^k n_i = n_1 + n_2 + \dots + n_k = N$$

n_1, n_2, \dots, n_k sono le frequenze delle modalità x_1, x_2, \dots, x_k

Frequenze relative e percentuali

- **frequenze relative o proporzioni**: si ottengono rapportando le frequenze assolute al totale delle unità, N . Indicheremo con f_1, f_2, \dots, f_k tali quantità, essendo

$$f_i = \frac{\text{frequenza della modalità } x_i}{\text{numero totale di osservazioni}} = \frac{n_i}{N}, i=1, 2, \dots, k.$$

- **frequenze percentuali** p_1, p_2, \dots, p_k : si ottengono moltiplicando per 100 le frequenze relative:

$$p_i = f_i * 100, \quad i=1, 2, \dots, k$$

$$\sum_{i=1}^k f_i = f_1 + f_2 + \dots + f_k = 1$$

$$\sum_{i=1}^k p_i = p_1 + p_2 + \dots + p_k = 100$$

Esempi di distribuzioni di frequenze assolute, relative e percentuali

| Genere | |
|---------------|----------|
| x_i | n_i |
| x_1 Femmine | 7 n_1 |
| x_2 Maschi | 13 n_2 |
| Totale | 20 N |

| Genere | |
|---------------|------------|
| x_i | f_i |
| x_1 Femmine | 0.35 f_1 |
| x_2 Maschi | 0.65 f_2 |
| Totale | 1 |

| p_i |
|------------|
| 35 p_1 |
| 65 p_2 |
| 100 |

| Disturbo | |
|-------------------|---------|
| x_i | n_i |
| x_1 Discalculia | 8 n_1 |
| x_2 Disgrafia | 4 n_2 |
| x_3 Dislessia | 8 n_3 |
| Totale | 20 |

| x_i | f_i | p_i |
|-------------------|-----------|----------|
| x_1 Discalculia | 0.4 f_1 | 40 p_1 |
| x_2 Disgrafia | 0.2 f_2 | 20 p_2 |
| x_3 Dislessia | 0.4 f_3 | 40 p_3 |
| Totale | 1 | 100 |

Esempio di distribuzione di frequenze assolute, relative e percentuali

Alunni con DSA per tipologia di disturbo-a.s. 2017/2018

| Tipologia di disturbo | <i>Frequenza assoluta</i> | <i>frequenza relativa</i> | <i>frequenza percentuale</i> |
|------------------------------|----------------------------------|----------------------------------|-------------------------------------|
| Dislessia | 177212 | 0,407 | 40,7 |
| Disgrafia | 79261 | 0,182 | 18,2 |
| Disortografia | 92134 | 0,212 | 21,2 |
| Discalculia | 86645 | 0,199 | 19,9 |
| TOTALE | 435252 | 1 | 100 |

Fonte Istat

Frequenze cumulate

Consideriamo una distribuzione di frequenze secondo un carattere a modalità ordinabili.

- Si chiamano **frequenze cumulate** le quantità

$$N_i = n_1 + n_2 + \dots + n_i, i = 1, 2, \dots, k.$$

Per ogni dato i , N_i rappresenta il numero delle unità del collettivo nelle quali il carattere X assume un valore non superiore a x_i

- Si chiamano **frequenze relative cumulate** i rapporti

$$F_i = \frac{\text{frequenza cumulata fino a } x_i}{\text{numero totale di osservazioni}} = \frac{N_i}{N}, i = 1, 2, \dots, k.$$

- **frequenze percentuali cumulate** : si ottengono moltiplicando per 100 le frequenze relative cumulate

Esempio di frequenze cumulate (1)

| ID | N.of previous offenses | education |
|----|------------------------|-----------------|
| 1 | 1 | high edu |
| 2 | 2 | low sec |
| 3 | 2 | primary or less |
| 4 | 10 | primary or less |
| 5 | 6 | high edu |
| 6 | 3 | low sec |
| 7 | 3 | primary or less |
| 8 | 5 | high edu |
| 9 | 4 | low sec |
| 10 | 8 | high edu |
| 11 | 7 | low sec |
| 12 | 4 | primary or less |
| 13 | 8 | primary or less |
| 14 | 4 | high edu |
| 15 | 6 | primary or less |
| 16 | 9 | upp sec |
| 17 | 9 | low sec |
| 18 | 3 | upp sec |
| 19 | 9 | high edu |

| N.of previous offenses committed | frequenza assoluta | frequenza cumulata | | |
|----------------------------------|--------------------|--------------------|-----------------|--------------------|
| | | <i>assoluta</i> | <i>relativa</i> | <i>percentuale</i> |
| 1 | 1 | 1 | 0.0526 | 5.26 |
| 2 | 2 | 3 | 0.1579 | 15.79 |
| 3 | 3 | 6 | 0.3158 | 31.58 |
| 4 | 3 | 9 | 0.4737 | 47.37 |
| 5 | 1 | 10 | 0.5263 | 52.63 |
| 6 | 2 | 12 | 0.6316 | 63.16 |
| 7 | 1 | 13 | 0.6842 | 68.42 |
| 8 | 2 | 15 | 0.7895 | 78.95 |
| 9 | 3 | 18 | 0.9474 | 94.74 |
| 10 | 1 | 19 | 1 | 100 |
| totale | 19 | | | |

Esempio di frequenze cumulate (2)

| ID | N.of previous offenses | education |
|----|------------------------|-----------------|
| 1 | 1 | high edu |
| 2 | 2 | low sec |
| 3 | 2 | primary or less |
| 4 | 10 | primary or less |
| 5 | 6 | high edu |
| 6 | 3 | low sec |
| 7 | 3 | primary or less |
| 8 | 5 | high edu |
| 9 | 4 | low sec |
| 10 | 8 | high edu |
| 11 | 7 | low sec |
| 12 | 4 | primary or less |
| 13 | 8 | primary or less |
| 14 | 4 | high edu |
| 15 | 6 | primary or less |
| 16 | 9 | upp sec |
| 17 | 9 | low sec |
| 18 | 3 | upp sec |
| 19 | 9 | high edu |

| Titolo di studio | frequenza assoluta | frequenza cumulata | | |
|------------------|--------------------|--------------------|-----------------|--------------------|
| | | <i>assoluta</i> | <i>relativa</i> | <i>percentuale</i> |
| primary or less | 6 | 6 | 0.3158 | 31.58 |
| low sec | 5 | 11 | 0.5789 | 57.89 |
| upp sec | 2 | 13 | 0.6842 | 68.42 |
| high edu | 6 | 19 | 1.0000 | 100.00 |
| totale | 19 | | | |

Distribuzioni doppie

- Considerando congiuntamente due colonne della matrice dei dati, l'insieme delle coppie di modalità dei due caratteri che così si osservano costituisce una **distribuzione doppia disaggregata**
- Le distribuzioni doppie di **frequenze** sono il risultato dello spoglio dei dati basato su una preliminare definizione delle modalità e delle eventuali classi per entrambi i caratteri.

Esempio di distribuzione doppia (1)

| SOGGETTO | GENERE | TIPO DI DISTURBO |
|----------|---------|------------------|
| 1 | MASCHIO | DISGRAFIA |
| 2 | FEMMINA | DISCALCULIA |
| 3 | FEMMINA | DISGRAFIA |
| 4 | MASCHIO | DISLESSIA |
| 5 | MASCHIO | DISLESSIA |
| 6 | FEMMINA | DISLESSIA |
| 7 | MASCHIO | DISGRAFIA |
| 8 | MASCHIO | DISCALCULIA |
| 9 | MASCHIO | DISCALCULIA |
| 10 | FEMMINA | DISGRAFIA |
| 11 | MASCHIO | DISCALCULIA |
| 12 | MASCHIO | DISCALCULIA |
| 13 | MASCHIO | DISLESSIA |
| 14 | MASCHIO | DISLESSIA |
| 15 | FEMMINA | DISCALCULIA |
| 16 | MASCHIO | DISLESSIA |
| 17 | FEMMINA | DISCALCULIA |
| 18 | FEMMINA | DISLESSIA |
| 19 | MASCHIO | DISCALCULIA |
| 20 | MASCHIO | DISLESSIA |

| Tipo di disturbo | Genere | | Totale |
|------------------|----------|-----------|-----------|
| | Femmina | Maschio | |
| Discalculia | 3 | 5 | 8 |
| Disgrafia | 2 | 2 | 4 |
| Dislessia | 2 | 6 | 8 |
| Totale | 7 | 13 | 20 |

Esempio di distribuzione doppia (2)

| SOGGETTO | GENERE | ORDINE DI SCUOLA |
|----------|---------|------------------|
| 1 | MASCHIO | INFANZIA |
| 2 | MASCHIO | SEC. I GRADO |
| 3 | MASCHIO | INFANZIA |
| 4 | MASCHIO | SEC. I GRADO |
| 5 | MASCHIO | SEC. I GRADO |
| 6 | MASCHIO | SEC. I GRADO |
| 7 | MASCHIO | PRIMARIA |
| 8 | MASCHIO | SEC. I GRADO |
| 9 | MASCHIO | INFANZIA |
| 10 | MASCHIO | INFANZIA |
| 11 | FEMMINA | INFANZIA |
| 12 | FEMMINA | INFANZIA |
| 13 | FEMMINA | INFANZIA |
| 14 | FEMMINA | INFANZIA |
| 15 | FEMMINA | PRIMARIA |
| 16 | FEMMINA | INFANZIA |
| 17 | FEMMINA | INFANZIA |
| 18 | FEMMINA | PRIMARIA |
| 19 | FEMMINA | INFANZIA |

| | GENERE | | |
|---------------|----------|-----------|-----------|
| ORDINE SCUOLA | FEMMINA | MASCHIO | Totale |
| INFANZIA | 7 | 4 | 11 |
| PRIMARIA | 2 | 1 | 3 |
| SEC. I GRADO | 0 | 5 | 5 |
| Totale | 9 | 10 | 19 |

Distribuzioni multiple

In generale si parla di:

- **distribuzione tripla**, se si considerano congiuntamente tre caratteri (tre colonne della matrice dei dati); **distribuzione quadrupla**, se si considerano congiuntamente quattro caratteri, e così via
- anche in questo caso, le distribuzioni si distinguono in disaggregate e di frequenze, a seconda che i dati si considerino allo stato grezzo o che si sia proceduto allo spoglio

Serie storiche

- Si ha una serie storica quando i dati statistici di interesse vengono associati a **modalità temporali**
- I dati statistici possono riguardare sia fenomeni di movimento (*flusso*) che fenomeni di stato (*stock*)

Esempio serie storica (fenomeno di movimento)

| | a.s. 1997/98 | | | a.s. 1998/99 | | | a.s. 1999/00 | | | a.s. 2000/01 | | |
|-------------------------|--------------|---------|-------------------|--------------|---------|-------------------|--------------|---------|-------------------|--------------|---------|-------------------|
| | Alunni | Classi | Alunni/ classi | Alunni | Classi | Alunni/ classi | Alunni | Classi | Alunni/ classi | Alunni | Classi | Alunni/ classi |
| Totale nazionale | 7.599.110 | 378.612 | 20,1 | 7.540.183 | 372.317 | 20,3 | 7.542.232 | 371.355 | 20,3 | 7.561.780 | 371.416 | 20,4 |

Fonte Istat: Alunni, classi e rapporto alunni/classi

Esempio serie storica (fenomeno di stato)

| Anno | 2005 | 2006 | 2007 | 2008 | 2009 |
|-----------------|-------------|-------------|-------------|-------------|-------------|
| N. di residenti | 58462 | 58752 | 59131 | 59616 | 60045 |

Fonte Istat Popolazione residente in Italia dal 2005 al 2009

Esempi di serie storiche

Serie Storica1–Alunni con cittadinanza non italiana (valori assoluti e percentuali) - A.S. 1983-84/2015-16

| Anni scolastici | Totale studenti con cittadinanza non italiana v.a. | Studenti con cittadinanza non italiana in rapporto alla popolazione scolastica totale % |
|-----------------|---|--|
| 1983/1984 | 6.104 | 0,06 |
| 1984/1985 | 6.468 | 0,06 |
| 1985/1986 | 7.050 | 0,07 |
| 1986/1987 | 7.424 | 0,07 |
| 1987/1988 | 8.967 | 0,09 |
| 1988/1989 | 11.791 | 0,12 |
| 1989/1990 | 13.668 | 0,14 |
| 1990/1991 | 18.794 | 0,19 |
| 1991/1992 | 25.756 | 0,27 |
| 1992/1993 | 30.546 | 0,32 |
| 1993/1994 | 37.478 | 0,41 |
| 1994/1995 | 43.876 | 0,47 |
| 1995/1996 | 50.322 | 0,56 |
| 1996/1997 | 57.595 | 0,66 |
| 1997/1998 | 70.657 | 0,81 |
| 1998/1999 | 85.522 | 1,1 |
| 1999/2000 | 119.679 | 1,47 |
| 2000/2001 | 147.406 | 1,70 |
| 2001/2002 | 196.414 | 2,20 |
| 2002/2003 | 239.808 | 2,70 |
| 2003/2004 | 307.141 | 3,50 |
| 2004/2005 | 370.803 | 4,20 |
| 2005/2006 | 431.211 | 4,80 |
| 2006/2007 | 501.420 | 5,60 |
| 2007/2008 | 574.133 | 6,41 |
| 2008/2009 | 629.360 | 7,03 |
| 2009/2010 | 673.800 | 7,52 |
| 2010/2011 | 710.263 | 7,86 |
| 2011/2012 | 755.939 | 8,42 |
| 2012/2013 | 786.630 | 8,87 |
| 2013/2014 | 803.053 | 9,00 |
| 2014/2015 | 814.208 | 9,20 |
| 2015/2016 | 814.851 | 9,23 |

Fonte



Serie territoriali

- Si ha una serie territoriale quando i dati vengono associati a modalità rappresentate da **entità territoriali**

Esempio serie territoriale

| Ripartizione territoriale | Alunni con DSA in % di alunni per area territoriale nella scuola primaria |
|----------------------------------|---|
| Italia Nord-Occidentale | 2.6 |
| Italia Nord-Orientale | 2.1 |
| Italia Centrale | 2.4 |
| Italia Meridionale | 1.1 |
| Italia | 2.0 |

Esempi di Serie territoriali

Tavola – Alunni con cittadinanza italiana e non italiana per regione – A.S. 2013/2014 e 2015/2016

| Regioni | Alunni con cittadinanza italiana | | | Alunni con cittadinanza non italiana | | | alunni italiani su alunni stranieri | |
|----------------|----------------------------------|-----------|------------|--------------------------------------|-----------|------------|---|-----------|
| | 2015/2016 | 2013/2014 | differenza | 2015/2016 | 2013/2014 | differenza | 2015/2016 | 2013/2014 |
| | Piemonte | 514.413 | 516.573 | -2.160 | 75.789 | 75.269 | 520 | 6,8 |
| Valle d'Aosta | 17.135 | 17.001 | 134 | 1.402 | 1.591 | -189 | 12,2 | 10,7 |
| Lombardia | 1.207.053 | 1.212.478 | -5.425 | 203.979 | 197.193 | 6.786 | 5,9 | 6,1 |
| Trentino A.A. | 144.234 | 146.931 | -2.697 | 18.433 | 17.709 | 724 | 7,8 | 8,3 |
| Veneto | 620.308 | 623.504 | -3.196 | 91.853 | 92.924 | -1.071 | 6,8 | 6,7 |
| Friuli V.G. | 142.156 | 142.964 | -808 | 18.960 | 19.018 | -58 | 7,5 | 7,5 |
| Liguria | 171.972 | 174.224 | -2.252 | 23.388 | 23.011 | 377 | 7,4 | 7,6 |
| Emilia Romagna | 520.875 | 516.075 | 4.800 | 96.213 | 93.434 | 2.779 | 5,4 | 5,5 |
| Toscana | 444.028 | 442.632 | 1.396 | 67.004 | 64.348 | 2.656 | 6,6 | 6,9 |
| Umbria | 106.173 | 106.086 | 87 | 16.945 | 17.341 | -396 | 6,3 | 6,1 |
| Marche | 196.061 | 196.215 | -154 | 25.439 | 26.543 | -1.104 | 7,7 | 7,4 |
| Lazio | 751.660 | 753.722 | -2.062 | 77.109 | 77.062 | 47 | 9,7 | 9,8 |
| Abruzzo | 171.138 | 174.272 | -3.134 | 13.260 | 13.245 | 15 | 12,9 | 13,2 |
| Molise | 40.304 | 41.926 | -1.622 | 1.450 | 1.486 | -36 | 27,8 | 28,2 |
| Campania | 985.200 | 1.011.845 | -26.645 | 22.492 | 21.779 | 713 | 43,8 | 46,5 |
| Puglia | 621.463 | 642.398 | -20.935 | 16.557 | 16.542 | 15 | 37,5 | 38,8 |
| Basilicata | 81.823 | 84.579 | -2.756 | 2.625 | 2.468 | 157 | 31,2 | 34,3 |
| Calabria | 290.345 | 299.853 | -9.508 | 12.580 | 12.921 | -341 | 23,1 | 23,2 |
| Sicilia | 769.405 | 791.680 | -22.275 | 24.319 | 24.128 | 191 | 31,6 | 32,8 |
| Sardegna | 216.296 | 222.217 | -5.921 | 5.054 | 5.041 | 13 | 42,8 | 44,1 |
| Italia | 8.012.042 | 8.117.175 | -105.133 | 814.851 | 803.053 | 11.798 | 9,8 | 10,1 |

Rappresentazioni grafiche

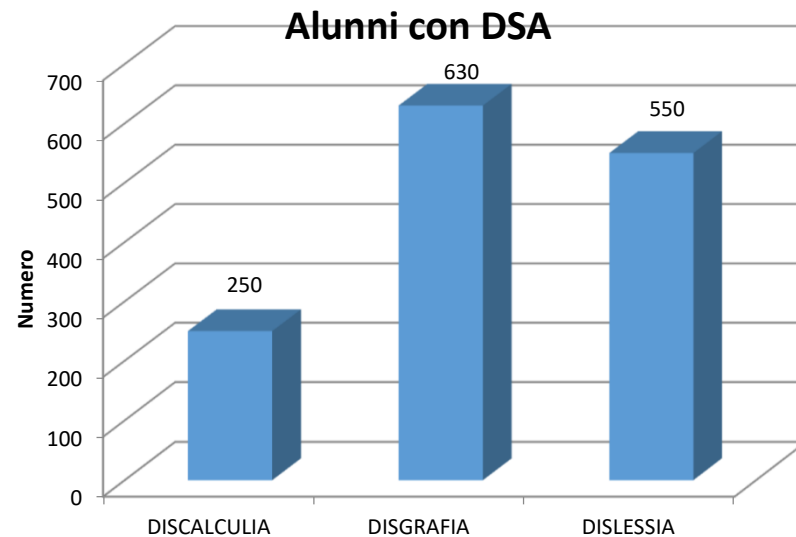
Per ogni distribuzione statistica, è possibile individuare la rappresentazione grafica più adatta, la soluzione, cioè, che concilia **correttezza metodologica** ed **efficacia informativa**

Scala nominale: grafico a nastri

Le distribuzioni di variabili nominali (serie sconnesse) vengono generalmente rappresentate con **grafici di tipo areale**, in cui alle modalità del carattere si fanno corrispondere figure geometriche (rettangoli, quadrati, settori circolari ecc.) con **aree proporzionali** alle grandezze da rappresentare. Le figure geometriche più spesso utilizzate sono i rettangoli (**nastri orizzontali o verticali**).

| TIPO DI DISTURBO | numero |
|--------------------|--------|
| DISCALCULIA | 250 |
| DISGRAFIA | 630 |
| DISLESSIA | 550 |
| Totale complessivo | 1430 |

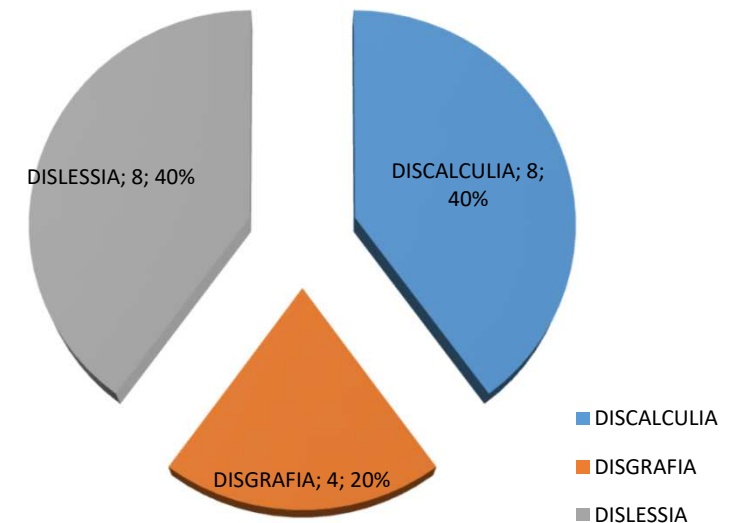
L'aspetto visivo del grafico rimane immutato se si utilizzano le frequenze relative o percentuali



Scala nominale: grafico a settori circolari o a torta

Una rappresentazione grafica alternativa, utilizzabile quando il numero delle modalità non è elevato, è il grafico a settori circolari, in cui le frequenze o le quantità associate alle varie modalità del carattere vengono rappresentate con le **aree dei settori circolari** in cui è suddiviso un cerchio

| Tipo di disturbo | frequenza assoluta | frequenza relativa | angolo |
|------------------|--------------------|--------------------|------------|
| x_i | n_i | f_i | α_i |
| DISLESSIA | 8 | 0.4 | 144 |
| DISGRAFIA | 4 | 0.2 | 72 |
| DISCALCULIA | 8 | 0.4 | 144 |
| Totale | 20 | 1 | 360 |



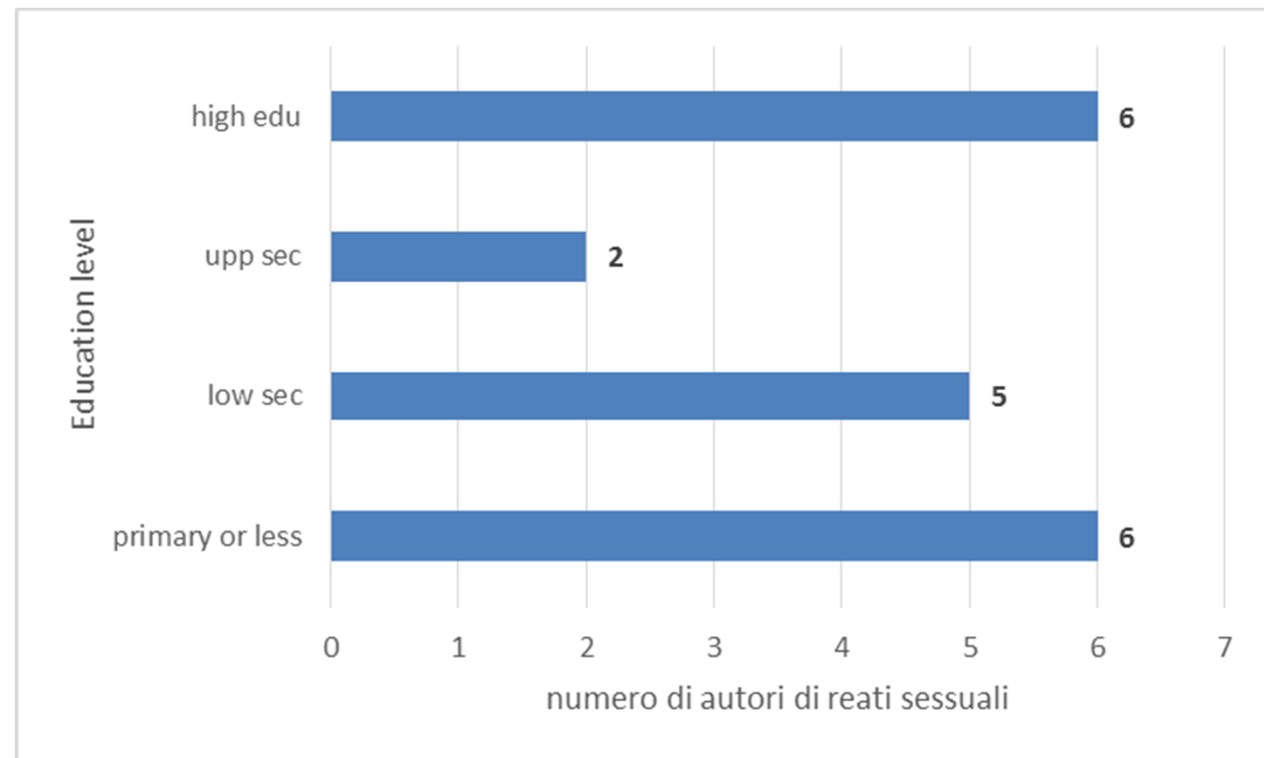
$$n_i : N = \alpha_i : 360^\circ$$

$$\alpha_i = \frac{n_i}{N} \cdot 360^\circ = f_i \cdot 360^\circ$$

Scala ordinale: grafico a nastri con modalità ordinate

Le distribuzioni di variabili ordinali vengono generalmente rappresentate con **grafici a nastri orizzontali o verticali** in cui le modalità devono essere ordinate

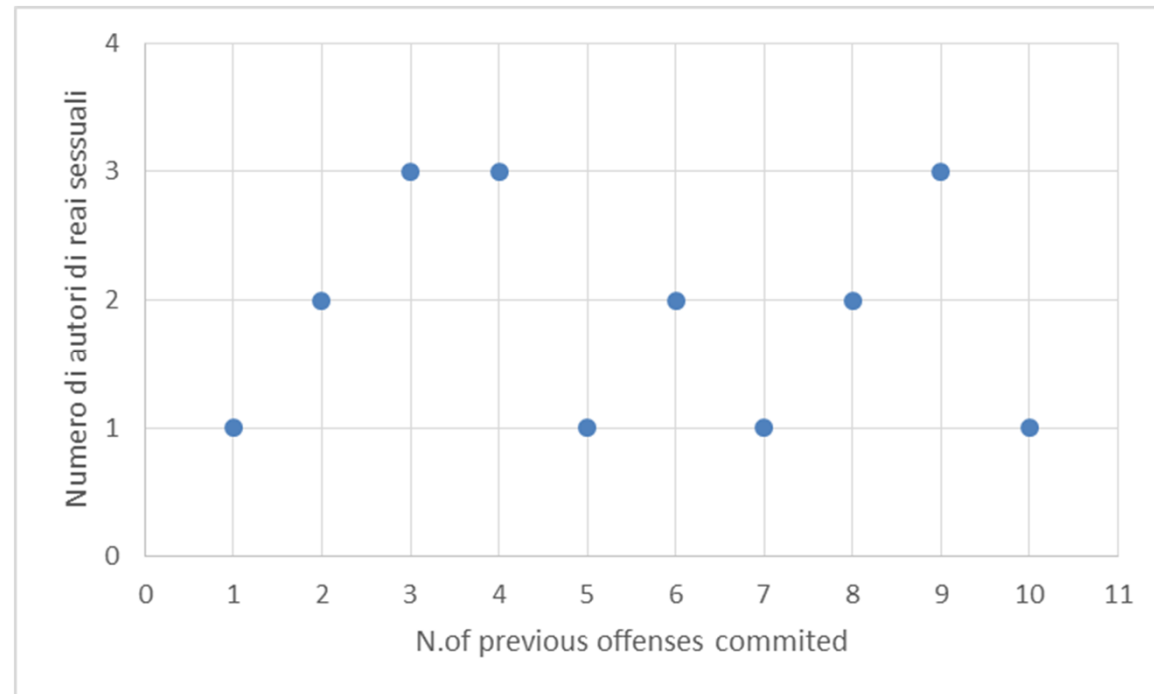
| Education level | frequenza assoluta |
|-----------------|--------------------|
| primary or less | 6 |
| low sec | 5 |
| upp sec | 2 |
| high edu | 6 |
| totale | 19 |



Caratteri quantitativi discreti: diagramma a punti

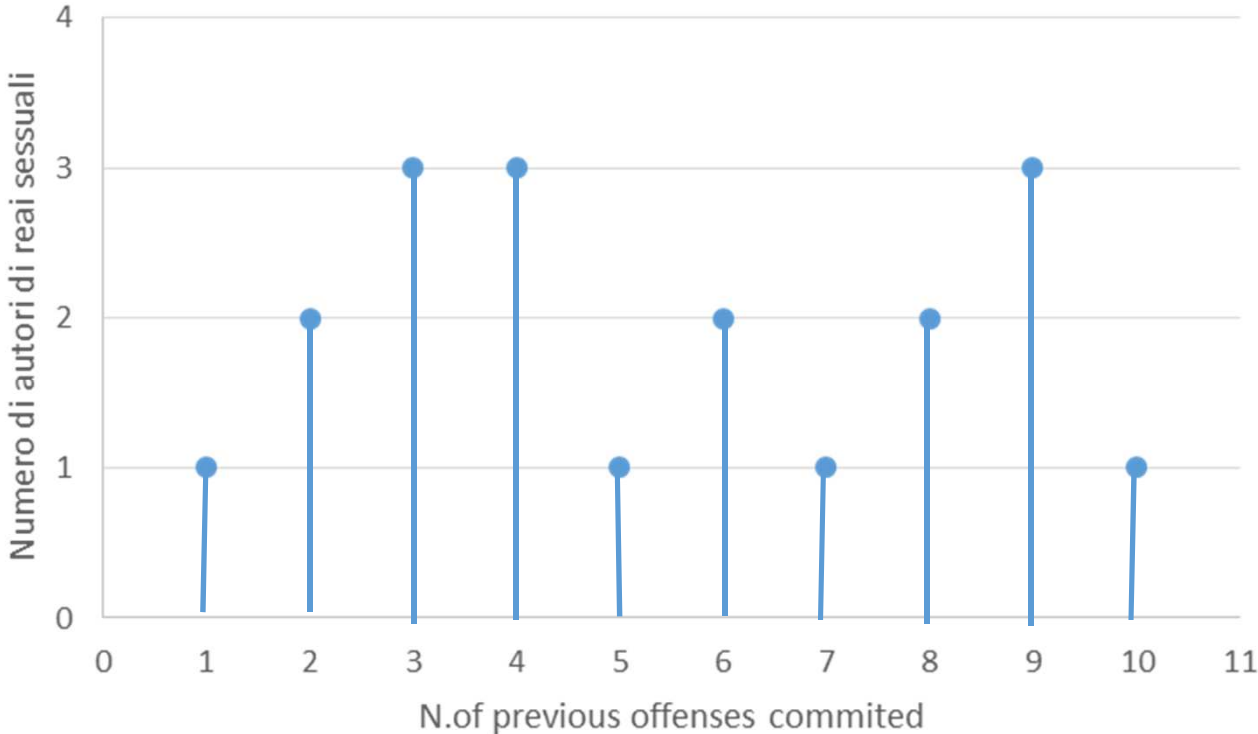
- La rappresentazione grafica più idonea per una distribuzione di frequenze secondo un carattere discreto è quella **cartesiana**
- Sull'asse delle ascisse vengono poste le modalità x_1, x_2, \dots, x_k , sull'asse delle ordinate le frequenze corrispondenti n_1, n_2, \dots, n_k
- La rappresentazione grafica va sotto il nome di **diagramma a punti**

| N.of previous offenses committed | frequenza assoluta |
|----------------------------------|--------------------|
| 1 | 1 |
| 2 | 2 |
| 3 | 3 |
| 4 | 3 |
| 5 | 1 |
| 6 | 2 |
| 7 | 1 |
| 8 | 2 |
| 9 | 3 |
| 10 | 1 |
| totale | 19 |



Caratteri quantitativi discreti: diagramma ad aste

| N.of previous offenses committed | frequenza assoluta |
|----------------------------------|---------------------------|
| 1 | 1 |
| 2 | 2 |
| 3 | 3 |
| 4 | 3 |
| 5 | 1 |
| 6 | 2 |
| 7 | 1 |
| 8 | 2 |
| 9 | 3 |
| 10 | 1 |
| totale | 19 |



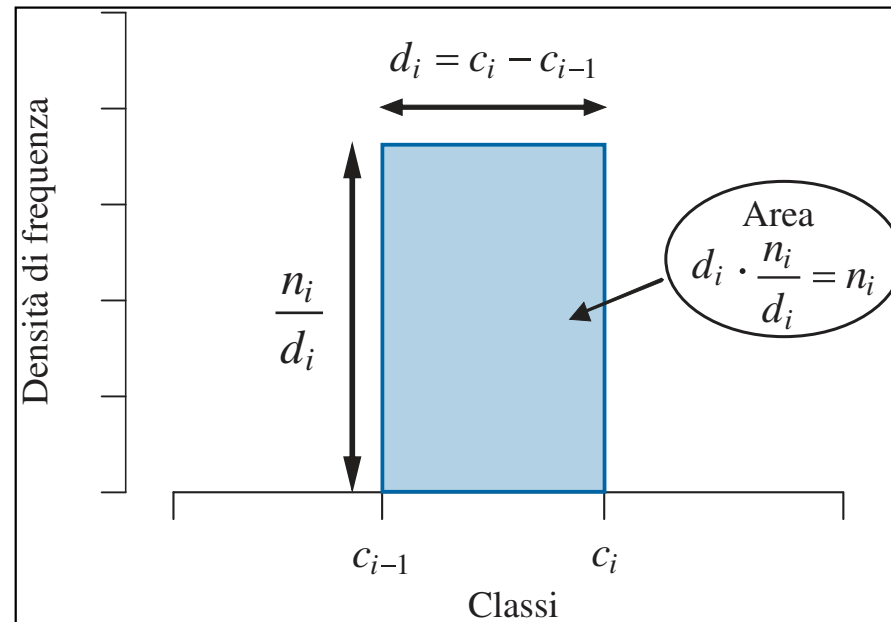
Caratteri quantitativi continui divisi in intervalli

La rappresentazione grafica più appropriata è l'istogramma in cui l'area del rettangolo rappresenta la frequenza assoluta o relativa della corrispondente classe.

L'istogramma si ottiene ponendo sull'asse delle ascisse gli estremi di classe c_0, c_1, \dots, c_k e disegnando per ogni classe (c_{i-1}, c_i) , $i = 1, 2, \dots, k$, un rettangolo avente per base il segmento dell'asse delle ascisse di estremi c_{i-1} e c_i e per altezza la densità di frequenza n_i/d_i , dove d_i è l'ampiezza di classe

| Classe ($c_{i-1}-c_i$) | Frequenza (n_i) |
|--------------------------|---------------------|
| $c_0 - c_1$ | n_1 |
| $c_1 - c_2$ | n_2 |
| \vdots | \vdots |
| $c_{k-1} - c_k$ | n_k |
| Totale | N |

Si chiama **distribuzione di frequenze** di un carattere X suddiviso in classi lo schema con cui si associa a ciascuna classe la rispettiva frequenza



Istogramma: esempio

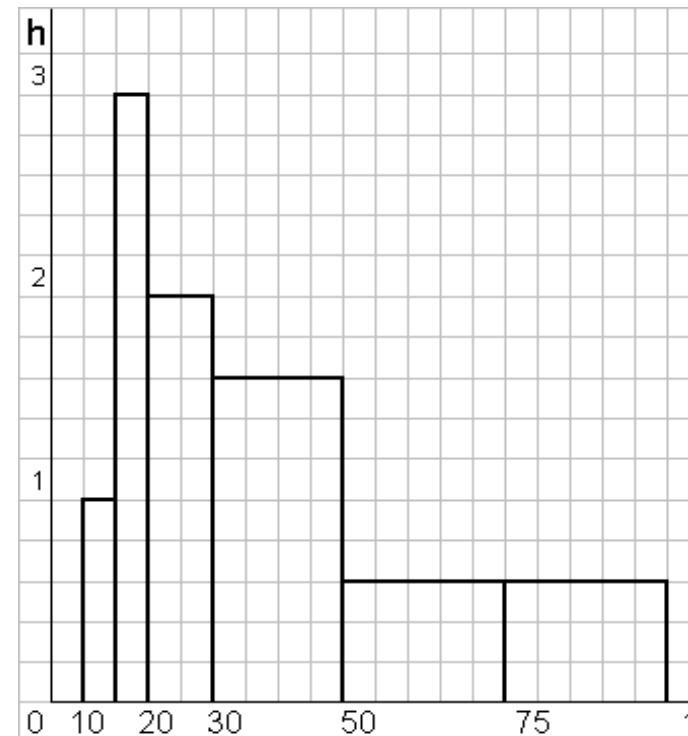
| Classi di peso (in Kg) | Frequenza assoluta | Ampiezza di classe | Densità di frequenza |
|---------------------------|-----------------------|-----------------------|-------------------------|
| $c_{i-1} - c_i$ | n_i | d_i | h_i |
| 10 -- 15 | 5 | 5 | 1 |
| 15 -- 20 | 15 | 5 | 3 |
| 20 -- 30 | 20 | 10 | 2 |
| 30 -- 50 | 30 | 20 | 1,5 |
| 50 -- 75 | 15 | 25 | 0,6 |
| 75 -- 100 | 15 | 25 | 0,6 |
| Totale | 100 | | |

Nell'istogramma si costruiscono dei rettangoli la cui area è proporzionale alla frequenza della classe.

La base è data dall'ampiezza di classe, l'altezza dalla densità di frequenza

$$d_i = c_{i-1} - c_i$$

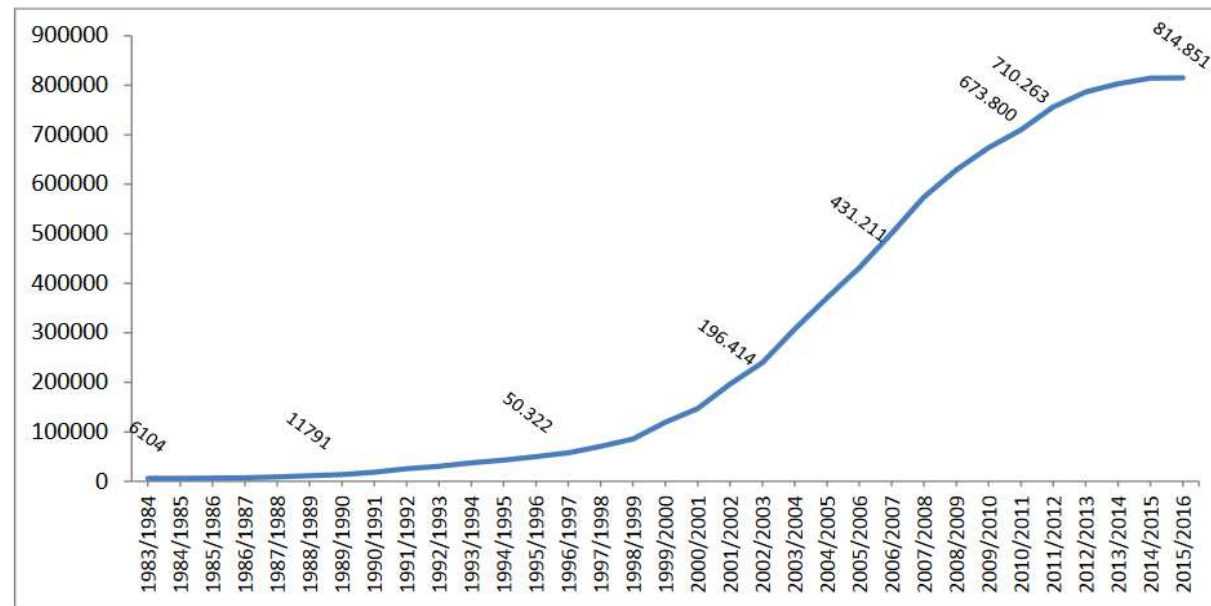
$$h_i = \frac{n_i}{d_i}$$



Grafici per serie storiche

- Per la rappresentazione grafica delle serie storiche si ricorre, generalmente, ai **diagrammi cartesiani**. Si pongono sull'asse delle ascisse i tempi e su quello delle ordinate le intensità associate: i conseguenti punti del piano cartesiano vengono poi uniti con **segmenti di retta** per facilitare la percezione visiva dell'andamento del fenomeno.

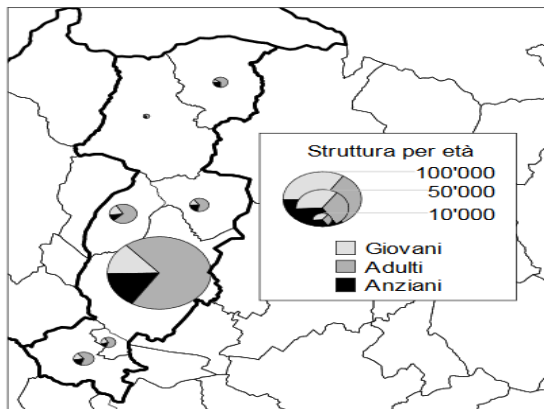
Grafico – Alunni con cittadinanza non italiana (valori assoluti) - AA.SS. 1983/1984 - 2015/2016



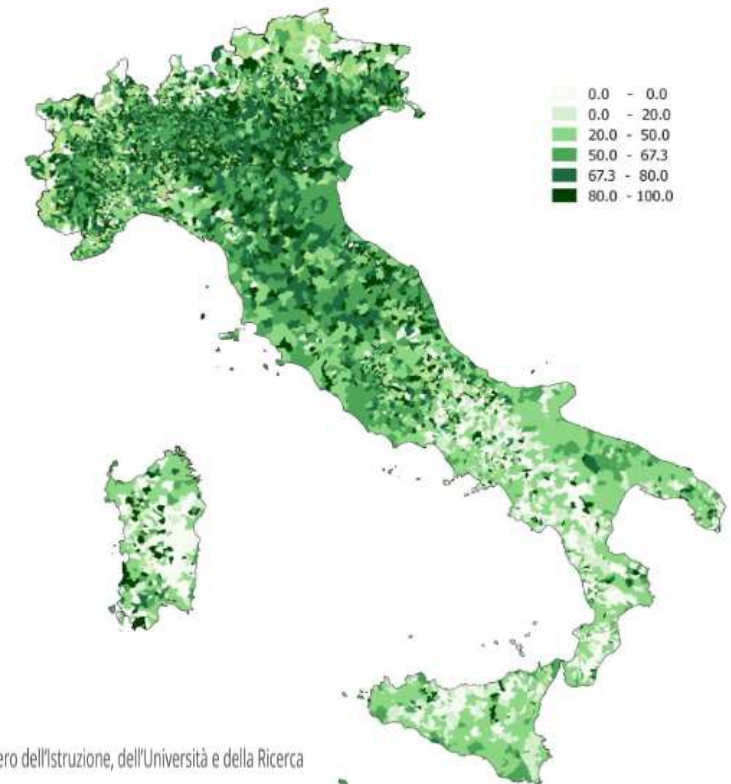
Fonte

Grafici per serie territoriali

- si rappresentano spesso mediante i **cartogrammi**: le ripartizioni territoriali sono individuate sull'appropriata **cartina geografica**; le intensità corrispondenti (frequenze o quantità) vengono rappresentate tramite **colori** o **tratteggi diversi**, il cui significato è specificato in apposite **legende**
- una variante è il **cartodiagramma**: a ogni ripartizione territoriale viene associato un grafico relativo a una distribuzione o ai livelli di uno o più fenomeni

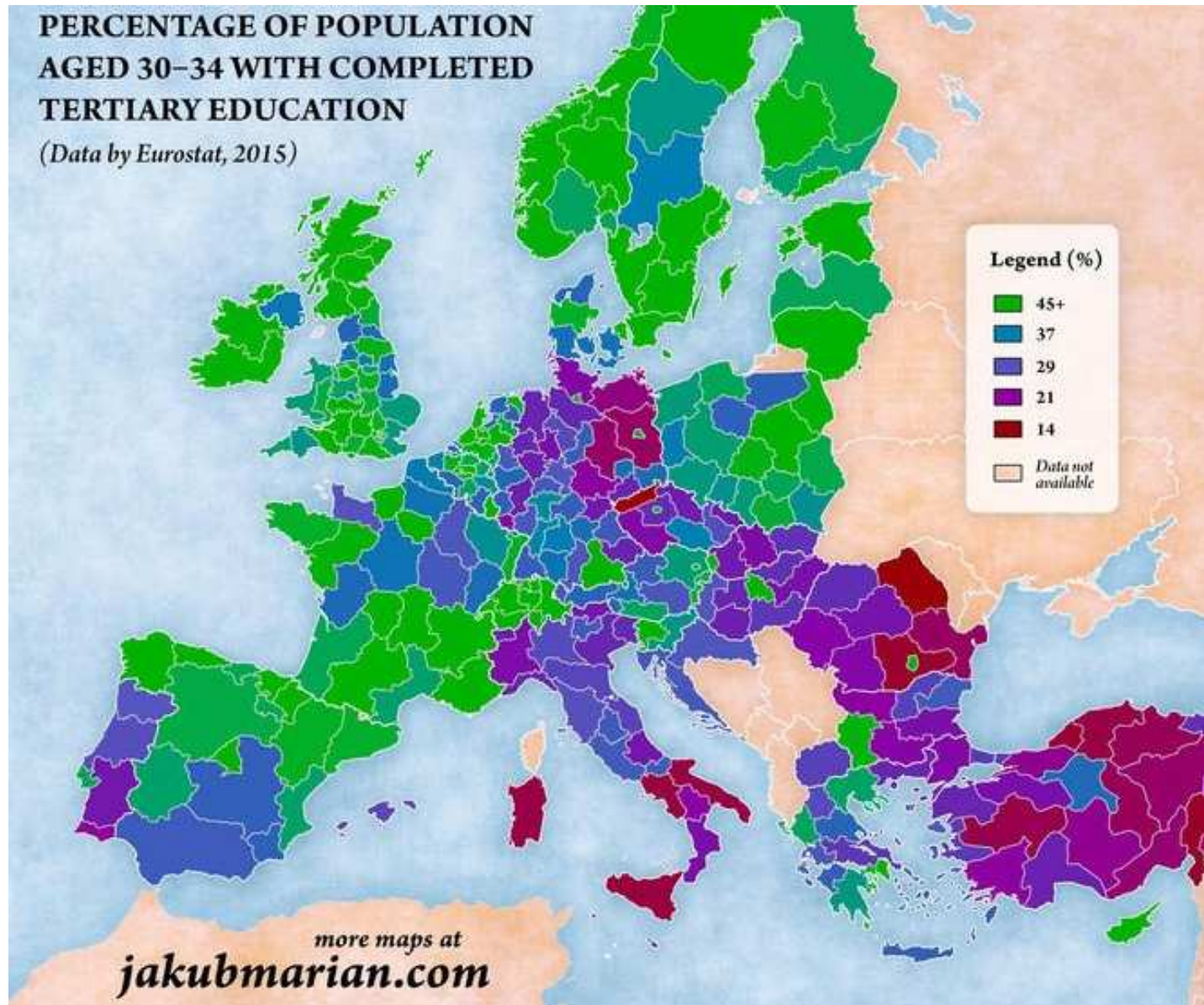


cartogramma relativo alla distribuzione comunale degli alunni con cittadinanza non italiana nati in Italia –A.S. 2015/2016



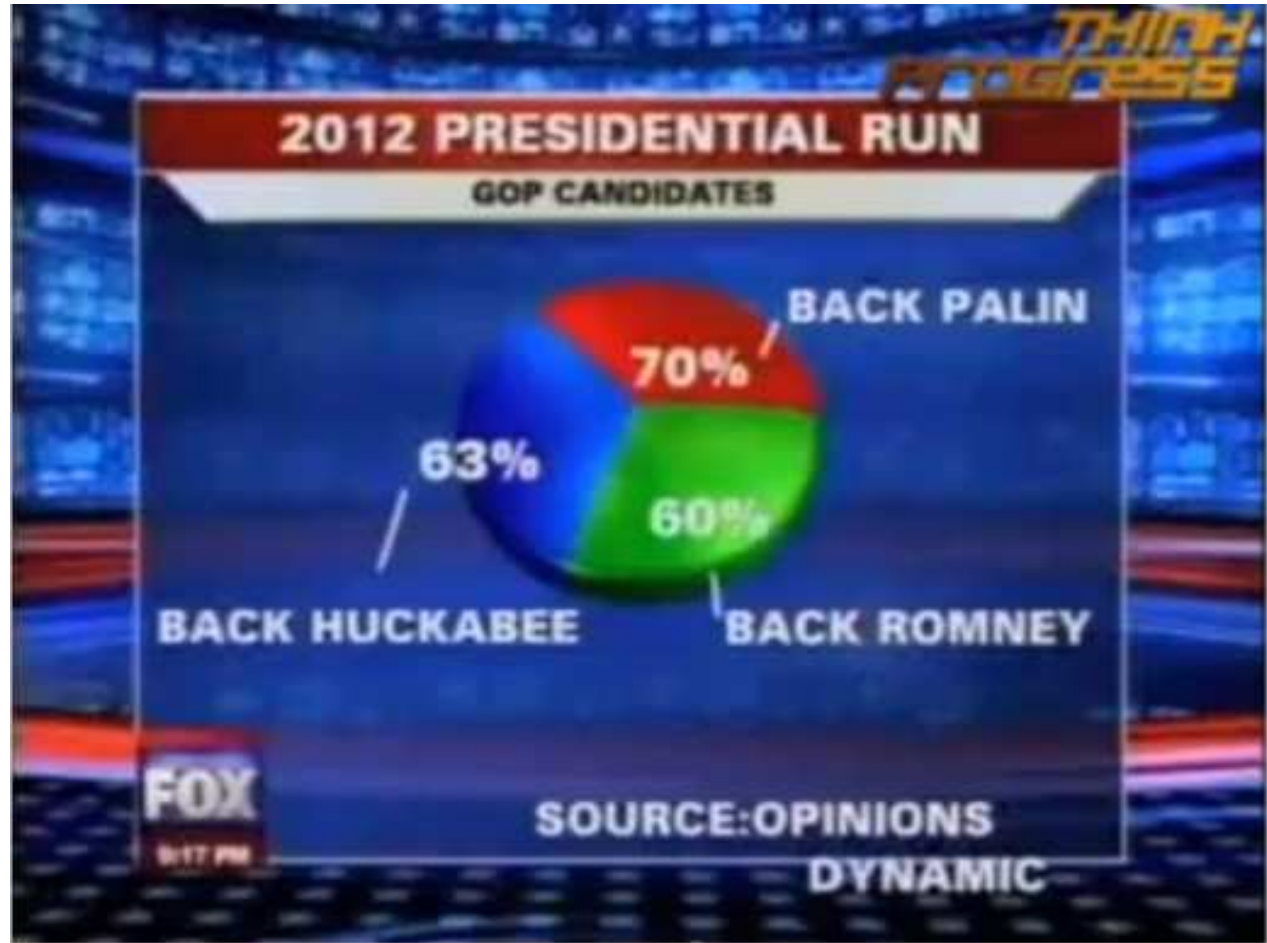
PERCENTAGE OF POPULATION AGED 30–34 WITH COMPLETED TERTIARY EDUCATION

(Data by Eurostat, 2015)



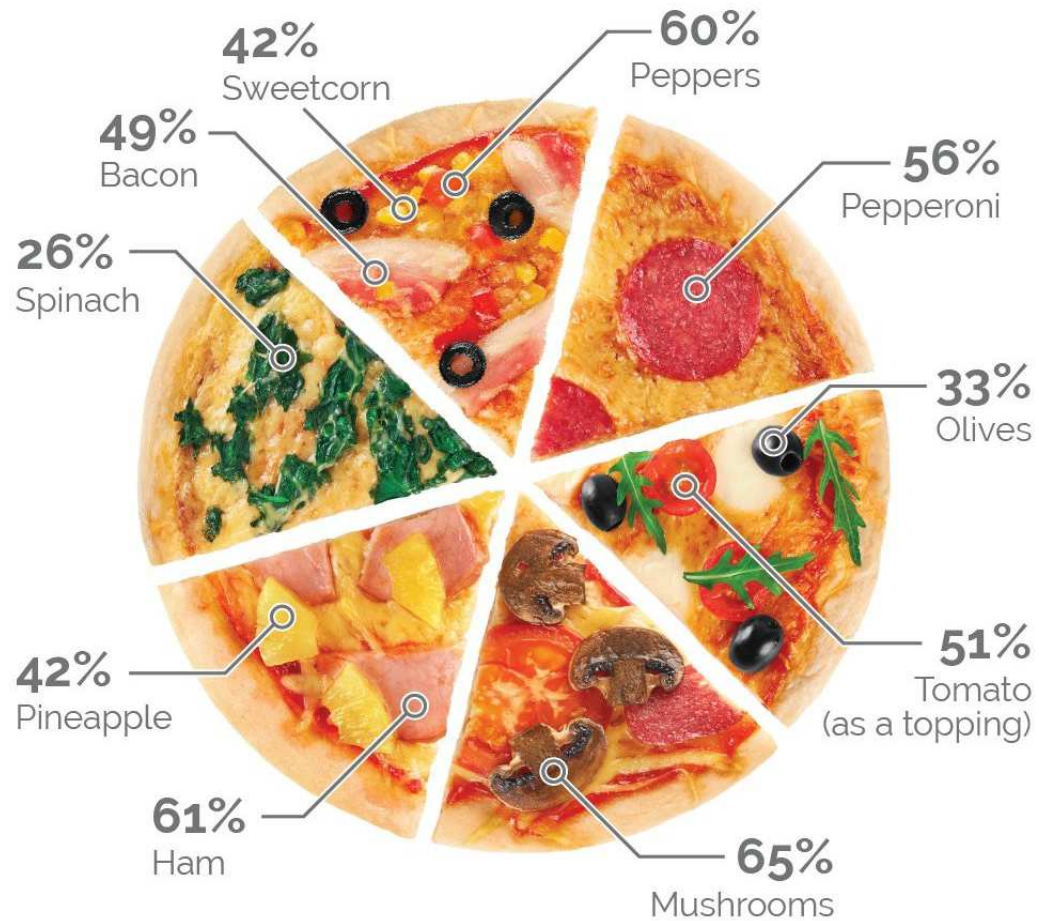
Errori nei grafici (1)

Numbers Don't Add Up



Mushroom is the UK's most liked pizza topping

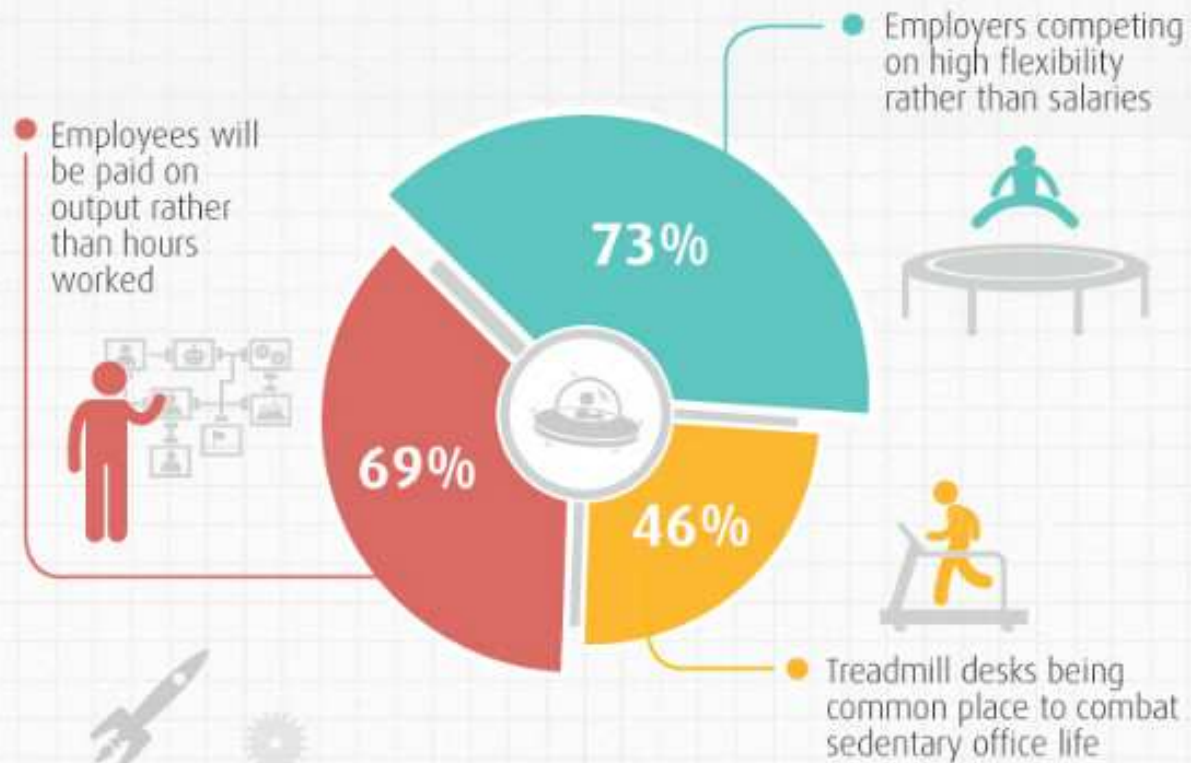
Generally speaking, which of the following toppings do you like on a pizza? Select as many as you like



Other items not depicted include: onions (62%), chicken (56%), beef (36%), chillies (31%), jalapeños (30%), pork (25%), tuna (22%), anchovies (18%). 2% of people say they only like Margherita pizzas

BIGGEST FUTURE CHANGES IN THE WORKPLACE?

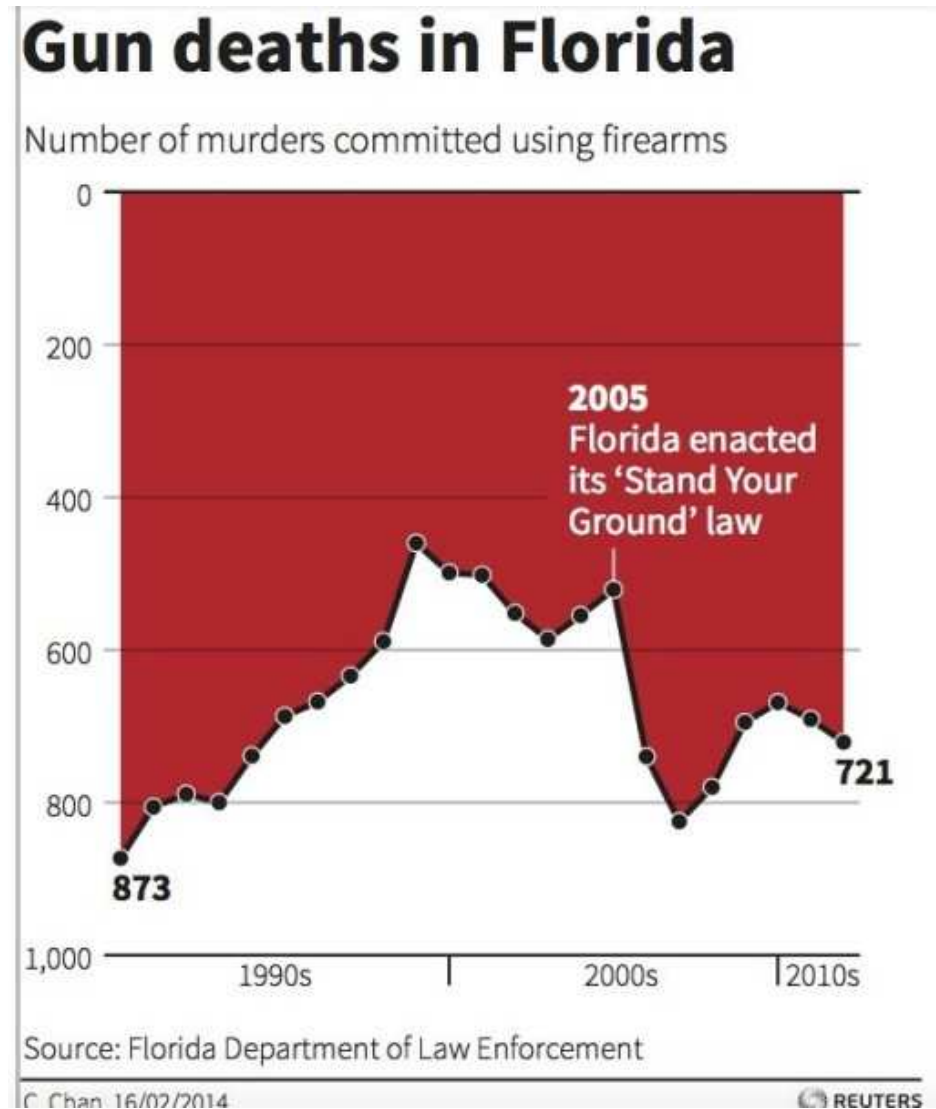
In a survey of HR professionals, changes forecasted in the workplace varied from:



Errori nei grafici (2)

Not Following Conventions

Quando guardi questo grafico cosa pensi sia successo fra il 2005 ed il 2012?



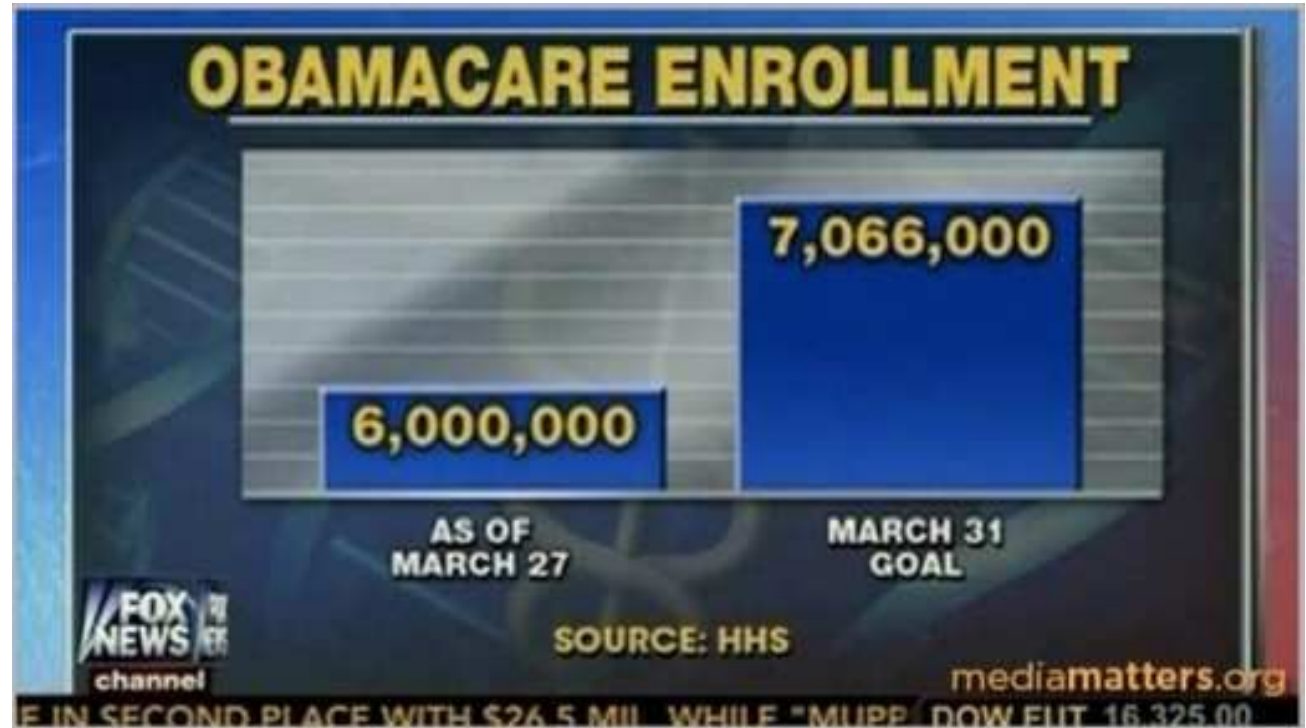


"He's right! When you look at it that way, it's not so bad!"

Errori nei grafici (3)

Cropped Axes

Quando guardi questo grafico pensi che l'obiettivo fissato per il 31 marzo sia facilmente raggiungibile?



E quando guardi questo grafico pensi che l'obiettivo fissato per il 31 marzo sia facilmente raggiungibile?



UNDER PRESIDENT OBAMA,
MORE STUDENTS ARE EARNING THEIR HIGH SCHOOL DIPLOMAS THAN EVER BEFORE

HIGH SCHOOL GRADUATION RATE



#LeadOnEducation

SOURCE: U.S. DEPARTMENT OF EDUCATION,
NATIONAL CENTER FOR EDUCATION STATISTICS

Errori nei grafici (3)

Improper Bubble Sizes

